AD-A118 141 FLORIDA UNIV GAINESVILLE F/G 12/1
ON THE RELATION BETWEEN STABLE MATRIX FRACTION FACTORIZATIONS A--ETC(U)
JAN 81 P P KHARGONEKAR, E D SONTAG DAAG29-80-C-0050
UNCLASSIFIED ARO-15077.13-MA NL

1 OF 1
AD A
118141

END
DATE
FILMED
09-82
DTIC

ARO 15077.13-MA 18

ON THE RELATION BETWEEN STABLE MATRIX FRACTION FACTORIZATIONS AND REGULABLE REALIZATIONS OF LINEAR SYSTEMS OVER RINGS

Pramod P. Khargonekar and Eduardo D. Sontag

Reprinted from IEEE Transactions on Automatic Control, Vol. AC-27, No. 3, June 1982
0018-9286/82/0600-0627$00.75 © 1982 IEEE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 15077.13-MA	2. GOVT ACCESSION NO. N/A AD-A118 144	3. RECIPIENT'S CATALOG NUMBER N/A
4. TITLE (and Subtitle) On the Relation Between Stable Matrix Fraction Factorizations and Regulable Realizations of Linear Systems Over Rings		5. TYPE OF REPORT & PERIOD COVERED Reprint
		6. PERFORMING ORG. REPORT NUMBER N/A
7. AUTHOR(s) Pramod P. Khargonekar Eduardo D. Sontag		8. CONTRACT OR GRANT NUMBER(s) DAAG29 80 C 0050
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Florida Gainesville, FL 32611		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS N/A
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P. O. Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE Jun 82
		13. NUMBER OF PAGES 12
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Submitted for announcement only.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

On the Relation Between Stable Matrix Fraction Factorizations and Regulable Realizations of Linear Systems Over Rings

PRAMOD P. KHARGONEKAR AND EDUARDO D. SONTAG

Abstract —Various types of transfer matrix factorizations are of interest when designing regulators for generalized types of linear systems (delay differential, 2-D, and families of systems). This paper studies the existence of stable and of stable proper factorizations, in the context of the theory of systems over rings. Factorability is related to stabilizability and detectability properties of realizations of the transfer matrix.

I. Introduction

THIS paper is motivated by recent research on the regulation ("servo problem") of certain classes of control systems which are "finite dimensional" and "linear" in a generalized sense. In contrast to the more standard linear *finite dimensional* case, linearity enters here in a more abstract sense, via the action of rings of operators or in terms of constraints on the quantities involved. For example, take a controlled delay equation like

$$(dx/dt)(t)=x(t)+3x(t-1)-u(t-2) \qquad (1.1)$$

whose natural state space is an infinite dimensional function space. This equation can be seen as a "finite dimensional" object if one introduces a ring of delay operators $\mathbf{R}[\theta]$, where $(\theta x)(t)=x(t-1)$, and then writes

$$(dx/dt)(t)=(1+3\theta)x(t)+(-\theta^2)u(t). \qquad (1.2)$$

This point of view suggests the use of methods from the usual theory in which coefficients of $x(t)$ and $u(t)$ are constant, but generalized to polynomial coefficients. In another example, when dealing with a discrete time system

$$x(t+1)=Fx(t)+Gu(t), \qquad (1.3)$$

one may want to restrict all control $u(t)$ and state values $x(t)$, as well as the entries of F and G to be integers; it is natural to model such restricted linear systems as systems over the ring of integers. In a variation of this last example, all quantities may be evaluated only modulo a fixed number r; for instance, $r=2^l$, $l=$ word length of a given computer. In this case one uses the ring of integers mod r to study a class of systems which are, in a sense, nonlinear. Yet another example of these generalized classes of systems is the situation in which one is interested in the study of parameterized classes of linear systems; this may be approached through the study of systems whose coefficients are functions of the parameters, with these functions having a specific structure (polynomial, analytic); the solution to a synthesis problem over the ring will provide a parameterized family of solutions to the corresponding problem for each system in the family. The literature on "systems over rings" is by now rather wide, and the reader is referred to the surveys Kamen [15] and Sontag [26], [28], and to the various papers on the subject in Byrnes and Martin [4], [5] for further motivations and examples.

Some of the generalized kinds of linear systems have been traditionally treated by other methods. This is especially so of various types of distributed systems, which can be studied via functional-analytic techniques or through the use of "frequency domain" (transfer matrix) design tools (see, e.g., Callier and Desoer [6]). We are interested here in the comparison to these latter methods, which involve an input/output approach in terms of various factorizations of transfer matrices (see, for instance, Youla *et al.* [31] and Desoer *et al.* [7]). The systems over rings approach is based on generalizations of state space techniques and in relation to the corresponding I/O maps. The main objective of this paper is to clarify the relationships between, on the one hand, the type of factorability assumptions made in the frequency domain approach, and on the other hand, properties of realizations and I/O maps over rings. It is not our purpose here to study problems of optimality, nor to characterize the class of all regulators achieving stability for a given plant; we concentrate solely on existence questions. The characterizations to be given in terms of stabilizability and detectability of realizations permit an intuitive understanding of many of the factorizations which employ rings of stable transfer functions. These factorizations can, in turn, be used in the study of other control-theoretic problems (output regulation, tracking, and disturbance rejection). Furthermore, some of the criteria given are useful in checking whether or not a given transfer matrix admits a factorization of the type needed, while

Manuscript received January 13, 1981; revised May 27, 1981. Paper recommended by E. W. Kamen, Chairman of the Linear Systems Committee. This work was supported by U.S. Army Research Grant DAAG 29-80-C0050 and U.S. Air Force Grants AFOSR-76-3034 ModD and AFOSR-80-0196.

P. P. Khargonekar is with the Center for Mathematical System Theory and the Department of Electrical Engineering, University of Florida, Gainesville, FL 32611.

E. D. Sontag is with the Department of Mathematics, Rutgers University, New Brunswick, NJ 08903.

0018-9286/82/0600-0627$00.75 ©1982 IEEE

other results exhibit the relations between existence of factorizations of different kinds.

It is important to note two major differences between our approach here and that in recent transfer matrix design methods. The first is that factorization and realization questions are considered here always in the context of a particular ring. For example, assume that one studies delay differential systems modeled over a ring $\mathbf{R}[\theta_1,\cdots,\theta_r]$, where θ_i is an a_i-second delay, for some rationally independent numbers a_i. Then all realizations, regulators, etc., are automatically systems over the same ring, i.e., delay systems whose delays are all integer combinations of the same basic length a_i's. In order to study more general realizations or regulators, one changes the base ring (for example, enlarges it to include more general operators from a suitable distribution ring as in Kamen [16]). Another major difference here is that causality is explicitly considered via a special role for a delay (in discrete-time) or differentiation (in continuous-time) operator z in the transfer matrices. This allows the realization of systems via difference or differential operators, depending on the interpretation of z.

II. Systems and Transfer Matrices

The results in this paper cannot be established without employing certain abstract concepts and results from commutative algebra. However, all of the results can be translated into "concrete" matrix theoretic terms for most base rings of system-theoretic interest, like a polynomial ring $K[\theta_1,\cdots,\theta_n]$ of polynomials in r variables with real ($K=\mathbf{R}$) or complex ($K=\mathbf{C}$) coefficients, or the ring of integers. Developing the theory only for such rings would be unnatural, since proofs would still be basically the same, but would have to be complemented with repeated use of the equivalences, e.g., "projective = free" (see below), which hold over these special rings. Furthermore, some rings for which the full "abstract" statements are needed are themselves of interest in system theory, e.g., certain residue rings. We shall therefore proceed in the following way. If the translation of an abstract concept does not follow—for the above-mentioned rings—from previously given translations, it will be given enclosed in slashes ($/\cdots/$) after the concept is introduced. Further, at various points the even more particular case of a one-variable polynomial ring $\mathbf{R}[\theta]$ will be used as an illustration, and in the last section an example is given which is also based on this ring.

In all that follows, R is an arbitrary, but fixed commutative ring (with identity). /Let R = integers or R = real or complex r-variable polynomial ring./ For the undefined algebraic terms, the reader should consult Bourbaki [1]. Examples of rings of system-theoretic interest are given in the references already mentioned; the last section will further restrict R. The term "module" will always mean *finitely generated* module over R or another ring if clear from the context, and *linear* will always mean linear with respect to the ring and modules in question. Composition of linear maps A, B will be denoted just by their juxtaposition AB. When modules are free, linear maps will be identified with their matrices in any fixed basis, with the same notation used for a map and its matrix. The integer $n(X)$ will denote the minimal possible cardinality of a set of generators for the module X. /All "modules" X of interest in definitions and statements can be assumed to be free, i.e., sets of column vectors $X=R^n$, where $n(X)=n$, endowed with the coordinatewise operations./ Two arbitrary integers m, p will be used to indicate number of input and output channels, respectively. When discussing any given system or transfer matrix, both the input- and output-value sets $U:=R^m$ and $Y:=R^p$ will be assumed fixed.

A *system* $\Sigma=(X,F,G,H,J)$ is given by a projective module X (the *state space*) and linear maps $F\colon X\to X$, $G\colon U\to X$, $H\colon X\to Y$, and $J\colon U\to Y$. When $J=0$, one has a *strictly causal* system; for such systems one drops J from the description.

Most rings that have appeared in the system theory literature are *projective-free*, i.e., projective modules over R are free. This includes polynomial rings in any number of variables with coefficients in a field, rings of continuous functions on contractible spaces, and principal ideal domains. Over such rings a system can then be thought of as a collection of matrices of appropriate sizes. Projective modules must still be used in developing the theory, however, since most constructions result in these.

/A *system* (R^n,F,G,H,J) is given by matrices F, G, H, J of sizes $n\times n$, $n\times m$, $p\times n$, and $p\times m$, respectively, *strictly causal* if $J=0$. For instance, over $R=\mathbf{R}[\theta]$ a "system" is really a family of classical linear systems $(F(\theta),G(\theta),H(\theta),J(\theta))$, parameterized by θ. These matrices can be seen as representing a discrete or continuous time linear system. In order to fix ideas, we shall follow this example through the paper as applied to continuous-time systems./

It is useful to introduce the *discrete-time interpretation* of a system. This corresponds to thinking of Σ as determining a set of equations

$$x(t+1)=Fx(t)+Gu(t), \tag{2.1a}$$

$$y(t)=Hx(t)+Ju(t) \tag{2.1b}$$

where $u(t)$, $x(t)$, and $y(t)$ denote input, state, and output values at integer times t. The use of (2.1) permits us to give many of the definitions, and to interpret the results in an intuitive way even if one is interested in other interpretations of the notion of system (e.g., delay-differential). Whatever definitions are given using the discrete time interpretation can be translated into algebraic properties of the maps defining a system, and in that sense they apply to all possible interpretations.

For example, it is natural to define a state to be *reachable* (from the origin) if it can be obtained as $x(T)$ for some $T>0$ when starting from $x(1)=0$ and solving (2.1a) with some sequence of inputs $u(1),\cdots,u(T-1)$. Let e_i be the ith element of the standard basis in U, $g_i:=Ge_i$, and $n=n(X)$. Then reachable states are precisely those in

$$\operatorname{span}\{F^jg_i,\ j=0,\cdots,n-1,\ i=1,\cdots,m\}. \tag{2.2}$$

A *system* is *reachable* if every state is reachable, i.e., if the span in (2.2) is all of X. Since reachable states are themselves controllable to zero (see below), reachability is equivalent to the complete controllability (any state can be driven to any other state) of the discrete time interpretation. One defines *observability* by the requirement that any two states be distinguishable by their input/output behavior; this corresponds with $n = n(X)$ to

$$\bigcap_{i=0}^{n-1} \ker HF^i = \{0\}. \tag{2.3}$$

(When R is an integral domain and X is free, (2.3) can be expressed as the usual Kalman observability condition.) A *canonical* system is one that is both reachable and observable. The *dual* of Σ is the system $\Sigma' := (X', F', H', G', J')$ where X' is the linear dual of X and F', etc. indicate transpose (adjoint) maps. Note that m, p are reversed for the dual system. A *coreachable* (or "strongly observable") system is one whose dual is reachable; this concept may be interpreted in terms of reachability of "observables" of the system. A *split* system is one that is both reachable and coreachable. The concepts discussed in this paragraph are by now relatively well-known; the references given before should be consulted for more details.

/*Reachability* means that block matrix $g(\Sigma) = [G, FG, \cdots, F^{n-1}G]$ has a right inverse over the ring R; *coreachability* means that $[H', F'H', \cdots, (F')^{n-1}H']$ has a *right inverse over* R; *split* if both hold. *Observability* corresponds to this last matrix being full-rank. Over $R = \mathbf{R}[\theta]$ reachability is equivalent to the gcd of the minors of $g(\Sigma)$ being a unit, or to the matrix $g(\Sigma)(\theta)$ being full-rank when evaluated at each complex number θ./

For purposes of regulation one needs much less than reachability and/or coreachability of a given system. In order to define the more general conditions, we first need some notion of stability and/or convergence. A purely algebraic way of introducing these is to postulate a set of "stable" polynomials to be fixed throughout the construction. This idea was used by Morse [24]. Our approach here follows Hautus and Sontag [14], but is generalized to the nonintegral domain case. A *Hurwitz set* S will be a multiplicative subset of admissible polynomials in $R[z]$ which contains at least one polynomial $z - a$ of degree one, and which is closed under associates. (A polynomial $p(z)$ is called *admissible* in Khargonekar [20] if there exists another polynomial $b(z)$ such that the product pb has a monic leading coefficient; for integral domains this means just that the leading coefficient of p is a unit in R. *Closed* under *associates* means that $a(z)p(z)$ belongs to S whenever $p(z)$ is in S and $a(z)$ is a unit in $R[z]$; for integral domains such an $a(z)$ is necessarily a constant.) The definitions and results to follow will always assume a Hurwitz set has been given.

/As a running example, for $R = \mathbf{R}[\theta]$, let S be the set of polynomials $p(\theta, z)$ in $R[z]$ which are monic in z and such that $p(v, s)$ is not zero for any v real and s with nonnegative real part. "Monic" can be taken to mean with leading coefficient in z equal to one, or simply a nonzero constant; this will not make any difference in what follows. Thus a polynomial in S represents a family of polynomials in $\mathbf{R}[z]$, parameterized by a scalar parameter θ, all of the same degree, and all Hurwitz in the usual (continuous-time) sense. The same type of example could, of course, be given for families parameterized by $r \neq 1$ variables.

Consider the ring $R((z^{-1}))$ of (formal) Laurent series over R. This is the set consisting of all formal sums

$$\Sigma a_t z^{-t}, \qquad a_t = 0 \qquad \text{for } t \leq t_0, \tag{2.4a}$$

where the ring operation is the usual multiplication. This ring was shown to be very useful in realization over rings by Wyman [30], and can be considered itself as a module over the subring $R[z]$ consisting of polynomials on nonnegative powers of z. Another subring of interest is the ring $R[(z)]$ of *rational* power series; this is the ring of fractions $T^{-1}R[z]$, where T is the set of polynomials in z which are admissible in the sense explained above. This ring of fractions will be identified with the subring (and $R[z]$-submodule) of $R((z^{-1}))$ obtained by long division into negative powers of z. For a Hurwitz set S one may consider the ring of fractions $S^{-1}R[z]$, which can also be seen as a subring of the Laurent series ring. The elements of this fraction ring will be called *stable* rational functions. For the discrete time interpretation it is useful to think of elements (2.4a) just as time functions

$$a(t) = a_t \tag{2.4b}$$

with support bounded to the left. The stable ones now will be interpreted as "(asymptotically) stable" sequences, and the notation

$$a(t) \text{"}\to\text{"} 0 \qquad (\text{as } t \to \infty) \tag{2.5}$$

will be used for these. (The quotation marks in the notation are included in order to emphasize that, for a particular Hurwitz set, these sequences may not converge in any reasonable sense. The "convergence" interpretation is very useful in guiding the proofs, as will be seen below, and the interest in applications is, of course, that in which either of these sequences indeed converge or they represent the coefficients of an expansion of a transform of a function which converges in the sense represented by the choice of S.) A *proper* (respectively, *strictly proper*) sequence (or series) will be one with $a(t) = 0$ for negative t (respectively, nonpositive t), i.e., power series in z^{-1} (respectively, with no constant term). For rational series and integral domains R properness corresponds to having a representation $p(z)/q(z)$ with $\deg(p) \leq \deg(q)$ ("$<$" for strictly proper). The subring of *proper stable* series will be denoted by $pr(R, S)$. The same notations will also be used for the set of Laurent series $M((z^{-1}))$ over a module M. This is a module over $R((z^{-1}))$ in the obvious way; the stable elements here are the elements of $S^{-1}M[z] = (S^{-1}R[z]) \otimes M$, and these form a module over $S^{-1}R[z]$. For example, in the case $M = R^n$, $S^{-1}R^n[z]$ is the set of all n vectors of stable series.

/For our running example of continuous-time scalar-parameter families, a series in $p(R, S)$ represents a family of transfer functions $p(\theta, z)/q(\theta, z)$ each of which is Hurwitz stable. The coefficients $a_i = a_i(\theta)$ give, then, for each θ the expansion of the Laplace transform of the corresponding stable transfer function./

An (asymptotically) *stable* system is one for which the characteristic polynomial $\det(zI - F)$ is in S. (Since state spaces are projective, characteristic polynomials are well-defined up to associates, by the method in Khargonekar [20].) /In our example, a stable system is, in fact, a family of continuous-time systems as before, such that each system in the family is stable./

A state x^* is (null-) *asymptotically controllable* (a.c.) if there exists an (infinite) input sequence $\{u(t), t \geq 1\}$ with $u(t)^* \to 0$ and such that the solution of (2.1a) with $x(1) = x^*$ also satisfies $x(t)^* \to 0$. Asymptotically controllable states form a submodule of X. A system is *asycontrollable* if every state is a.c. A system is *detectable* if its dual is asycontrollable. Note again that in the abstract setup there is no reason that these controllability concepts really should correspond to any such concrete notion, although of course this will be the case in the applications of interest (i.e., choice of Hurwitz set).

/For our example, a system (i.e., family of continuous-time systems) is asycontrollable (respectively, detectable) iff each member of the family is asycontrollable ("stabilizable" in the usual literature) (respectively, detectable); this is discussed in detail in Section VI. For nonscalar (real) parameters it is still an open problem whether or not asycontrollability of the family is equivalent to each system being stabilizable, but the above definition is still equivalent to a number of "spectral" types of conditions (and for complex parameters the equivalence always holds); this is discussed in Hautus and Sontag [14]./

A *rational* matrix or *transfer matrix* (with m inputs and p outputs) $W = W(z)$ is a $p \times m$ matrix whose entries are in $R[(z)]$. A [*strictly*] *proper* transfer matrix is one whose entires are all (strictly) proper. Any system Σ gives rise to a corresponding proper transfer matrix $W = W(\Sigma)$ defined by

$$W(z) = H(zI - F)^{-1}G + J. \tag{2.6}$$

When $J = 0$, W is strictly proper. Conversely, a (strictly) proper W always admits a *realization*, i.e., a (strictly causal) Σ satisfying (2.6). In fact, it is known from the realization theory over rings how to construct for any given W a *canonical* realization; this realization does not appear to be that useful for regulation questions, however, because the corresponding state space is, in general, not projective (except for special rings, like those in the last section). Finally, we define a transfer matrix to be *stable* if each of its entries is stable.

As remarked in the introduction, there has been interest lately in various types of factorizations of transfer matrices. One considers a

right factorization

$$W = PQ^{-1} \tag{2.7}$$

or a

left factorization

$$W = Q^{-1}P, \tag{2.8}$$

with P and Q either polynomial or rational matrices. (For the factorizations to be well-defined, we assume in the polynomial case that Q is admissible, i.e., that $\det Q$ is admissible in the sense defined before; in the rational case we assume that Q can be written as $q^{-1}Q^*$, with the polynomial q monic and the polynomial matrix Q^* admissible.) We are interested here in factorizations that satisfy a *Bezout condition*: for P and Q as in (2.7), this means that there exist matrices A and B such that

$$AP + BQ = I; \tag{2.9}$$

for P and Q as in (2.8) one wants A and B with

$$PA + BQ = I. \tag{2.10}$$

(In (2.9), B is square m by m, and A is m by p; dually for (2.10).) Depending on the type of matrices A and B allowed in the above, as well as on the allowed P and Q, one may then classify factorizations as: 1) *polynomial*, 2) *stable*, and 3) *proper stable*, meaning that *all* the matrices appearing—A, B, P, Q—must be of the corresponding type.

The main goal of this paper is to relate the various types of factorizations with the existence of realizations of different kinds. One of these relations is already known as follows.

Theorem 2.11 [20]: The following statements are equivalent for any strictly proper transfer matrix W:

1) W admits a polynomial right factorization;
2) W admits a polynomial left factorization;
3) W admits a split realization;
4) the canonical realization $\Sigma(W)$ is (projective) and split.

/In 4), read "the canonical realization exists and is split."

In Emre and Khargonekar [11] it is proved that a (free) split system can be regulated in much the same way as in the classical (linear finite-dimensional) case. Roughly, it is possible to achieve arbitrary dynamics both for the regulated input/output behavior and for the remaining (observer) modes of the closed-loop (plant/regulator) dynamics. Moreover, the split condition can be checked in various ways directly from W (see Sontag [26], [27] and Khargonekar [20]). If the system is just reachable and detectable, the same paper shows how essentially arbitrary dynamics can be achieved for the regulated I/O behavior while keeping the observer modes stable. Finally, Emre [10] proved that if the system is (only) asycontrollable and detectable, then it admits a stabilizing compensator. These results provide a strong motivation for the study of the existence of split, reachable/detectable, and asycontrollable/detectable realizations.

The existence of Bezout factorizations would appear to be rather restrictive when working over rings. A recent result of Lee and Olbrot [22], however, established the genericity of reachability, over R = polynomial ring in r variables, when the number of inputs m exceeds r. This result could be, in principle, extended to one about genericity of the split condition, although the form that the precise statement would take is not yet clear. This result would imply—via (2.4) and the results below—the genericity of the various types of Bezout factorability when dealing with systems with enough input and output channels. For "few" channels, the split condition—i.e., polynomial factorability—is, of course, too strong. For instance, the (scalar) transfer function $p(\theta, z)/q(\theta, z)$ defined over $R = \mathbf{R}[\theta]$ splits if and only if the plane complex curves determined by p, q do not intersect. The other Bezout conditions are not, in general, as strong however.

/Take, for instance, the case of real scalar families of systems mentioned in our example. There, stable factorability is equivalent to p, q having no common zeros with θ real and $\mathrm{Re}\, s \geq 0$, i.e., no unstable pole/zero cancellations for any member of the family of systems./

III. Polynomial Matrix Interpretations

The definitions of reachability and asycontrollability (and their duals) will be reformulated in a less intuitive, but more useful way in this section. Note first that the maps defining a system can be extended in the obvious (pointwise) way to the spaces $U((z^{-1}))$, $Y((z^{-1}))$, and $X((z^{-1}))$; these extensions are linear over $R((z^{-1}))$ (and all its subrings), and they will be denoted in the same way as the original maps. Consider now a fixed system Σ and Hurwitz set S.

A state $x^\wedge$ is reachable iff there exist polynomials $x(z)$ and $u(z)$ such that $(zI - F)x(z) + Gu(z) = x^\wedge$. This is just another way of saying that there is a finite input sequence $-u(t)$, zero for positive t, which drives the state 0 (at some time $t < 0$) to $x(1) = x^\wedge$. Let

$$[zI - F, G]: X[z] \oplus U[z] \to X[z] \tag{3.1}$$

denote the map that sends a pair of polynomials (x, u) into $(zI - F)x + Gu$, thought of as an $R[z]$-module map. The above then says that reachability is equivalent to X being contained in the image of (3.1). Since X generates the projective $R[z]$-module $X[z]$, we have the following.

Lemma 3.2: The following statements are equivalent for any system Σ:

1) Σ is reachable;

2) $[zI - F, G]$ is onto;

3) there exist linear maps $M(z)$: $X[z] \to X[z]$ and $N[z]$: $X[z] \to U[z]$ with

$$[zI - F, G]\begin{bmatrix} M(z) \\ N(z) \end{bmatrix} = I \qquad (\text{in } X[z]).$$

/Reachability is equivalent to right invertibility of $[zI - F, G]$ over $R[z]$./

Consider again the discrete time interpretation of Σ. We claim that any reachable state $x^\wedge$ is controllable to zero in finite time $n = n(X)$. More precisely, there exist sequences $x(t)$ and $u(t)$ satisfying (2.1a) and such that $x(1) = x^\wedge$, $x(t) = 0$ for $t > n$, and $u(t) = 0$ if $t \leq 0$ or $t \geq n$. This is proved in the same way as for finite dimensional systems over fields: by the Cayley–Hamilton theorem (which is valid over any commutative ring), one can write

$$F^n = \sum_{i=0}^{n-1} a_i F^i. \tag{3.3a}$$

so for $x^\wedge$ in the span (2.2) one has that

$$F^n x^\wedge = -\sum_{i=0}^{n-1} F^{n-i-1} G u_i \tag{3.3b}$$

for suitable u_i in U. Thus the sequence

$$u(t) := u_{t-1}, \qquad 1 \leq t \leq n \tag{3.4}$$

(and 0 otherwise) results from $x(1) = x^\wedge$ in $x(t) = 0$ for $t > n$. Equivalently, there exist polynomials $x(z)$ and $u(z)$ of degree strictly less than n such that

$$(zI - F)(z^{-n}x(z)) + G(z^{-n}u(z)) = x^\wedge. \tag{3.5}$$

Here the map in (3.1) is seen as a map between rational or Laurent series. We would like to conclude from here that $x^\wedge$ is a.c.; however, a finite sequence is not necessarily stable (i.e., z may not be in the Hurwitz set S being considered). But the following may be used. Let $x^\wedge$ be reachable, and consider the *new* system defined over the same state space but with $A := F + aI$ instead of the original F (here a is such that $z - a$ is in S, and I is the identity map in X). The state $x^\wedge$ is still reachable in this new system (just note that the generators of (2.2) using A are in the span of the original ones and vice versa). So an equation like (3.5) holds, with A in the place of F. Applying the substitution in $R[(z)]$ (ring homomorphism) $z \to z - a$, there results an equation

$$(zI - F)x^*(z) + Gu^*(z) = x^\wedge \tag{3.6}$$

where $x^*(z) = x(z-a)/(z-a)^n$ and $u^*(z) = u(z-a)/(z-a)^n$ are both stable and strictly proper. We conclude as follows.

Lemma 3.7: Reachable states are asymptotically controllable.

Let A, B be (R-) modules. Elements of $S^{-1}(\mathrm{Hom}_R(A, B))[z]$ can be naturally seen as $S^{-1}R[z]$-module (or equivalently, $R[z]$-module) maps from $S^{-1}A[z]$ into $S^{-1}B[z]$; when B is finitely presented (e.g., if projective) every such map can be represented in this form (see Bourbaki [1, sect. II.2.7]). A (strictly) *proper* map from $S^{-1}A[z]$ into $S^{-1}B[z]$ will be, by definition, one which is a (strictly) proper element of $S^{-1}(\mathrm{Hom}_R(A, B))[z]$ under this identification. A proper $R[z]$-map from $S^{-1}A[z]$ into $S^{-1}B[z]$ can also be identified with an $R[z]$-map from $\mathrm{pr}(R, S) \otimes A$ into $\mathrm{pr}(R, S) \otimes B$. /For all modules consisting of column vectors, identify matrices of rational functions with (rational) series whose coefficients are matrices./

Now we can obtain the following desired characterizations.

Proposition 3.8: The following statements are equivalent for any system Σ:

1) Σ is asycontrollable;

2) $[zI-F, G]$: $S^{-1}X[z] \oplus S^{-1}U[z] \to S^{-1}X[z]$ is onto;

3) there exist linear maps $M(z)$: $S^{-1}X[z] \to S^{-1}X[z]$ and $N(z)$: $S^{-1}X[z] \to S^{-1}U[z]$ with

$$[zI-F, G]\begin{bmatrix} M(z) \\ N(z) \end{bmatrix} = I \qquad (\text{in } S^{-1}X[z]);$$

4) there exist M, N *proper* as above.

Asycontrollability is equivalent to $[zI-F, G]$ having a stable (proper stable) right inverse. /

Proof: Since X is projective, $S^{-1}X[z]$ is projective over $S^{-1}R[z]$, so 2) and 3) are equivalent. Assume now that 1) holds. Then any x^* in X is in the image of the map in 2). Since X spans this image, 2) holds. We shall prove now that 4) implies 1) and that 3) implies 4).

Let $M(z)$ and $N(z)$ be as in 4), and pick any x^* in X. Let $x^*(z) := M(z)x^*$, $u^*(z) := N(z)x^*$. These are proper, stable, and satisfy (3.6). Thus if $x^*(z) = x_0 + x_1 z^{-1} + \cdots$, $u^*(z) = u_0 + u_1 z^{-1} + \cdots$, comparing powers of z in (3.6), one concludes that $x_0 = 0$ and that $x_1 + Gu_0 = x^*$. Since stable sequences form a submodule, we conclude that $u^*(z) - u_0$ is also stable, and controls x_1 asymptotically to zero. But Gu_0 is reachable, so by Lemma 3.7 it is also a.c.; thus x^* is a sum of a.c. states, and is a.c. itself. So 1) holds.

Finally, we prove that 3) implies 4). Consider the map

$$\left[(z-a)^{-1}(zI-F), G\right]:$$
$$\mathrm{pr}(R,S) \oplus (X+U) \to \mathrm{pr}(R,S) \otimes X. \quad (3.9)$$

We shall prove that this map is onto for the shown domain and codomain. It is enough for this to check that (3.9) is onto when reducing modulo every maximal ideal of $\mathrm{pr}(R,S)$ (i.e., tensoring by the possible residue fields of the latter). These evaluations are of two types: 1) those in which $(z-a)^{-1}$ reduces to zero, and 2) those that extend to the ring $S^{-1}R[z]$. This is because the latter can also be seen as the ring of fractions of $\mathrm{pr}(R,S)$ with respect to the multiplicative set generated by $(z-a)^{-1}$. The evaluations of type 1) (intuitively, "at $z=\infty$") give the identity for the first block in (3.9), so the map is indeed onto at these ideals (see Hautus and Sontag [14] for more details). The evaluations of type 2) can be seen as reductions of those appearing in the hypothesis 3), so they are also onto. Thus (3.9) is indeed onto. There exist, then, proper $M(z)$, $N(z)$ such that

$$(zI-F)M_1(z) + GN(z) = I \qquad (3.10)$$

where $M_1(z) = (z-a)^{-1}M(z)$ is also proper. (In fact, even if only $N(z)$ were known to be proper, the equation

$$M_1(z) = (zI-F)^{-1}(I - GN(z)) \qquad (3.11)$$

shows that M_1 is proper.) #

The statements in Lemma 3.2 and Proposition 3.8 can all be dualized into left invertibility conditions

$$\begin{bmatrix} M(z) & N(z) \end{bmatrix}\begin{bmatrix} zI-F \\ H \end{bmatrix} = I \qquad (3.12)$$

for coreachability and detectability. This equation can be again understood over the original spaces without dualizing because all modules in the equations are projective and thus also reflexive (double dual equal to the original module).

The above results are especially interesting when X is a free module. In that case, the conditions become just left or right invertibility with respect to stable or proper stable matrices. Note also that Proposition 3.8, statement 4) is the definition of stabilizability given in Hautus and Sontag [14], while Proposition 3.8, statement 3) is the one used by Emre [10]. All these definitions are then equivalent to the one given in Section II.

The definitions given here allow one to prove a number of facts in the "natural" way. Take for instance the lemma after the following definition.

Definition 3.13: A *regulator* for the strictly causal system $\Sigma = (X, F, G, H)$ is a system $\Sigma^{(r)} = (X^{(r)}, A, B, C, D)$ with input value set $U \oplus Y$ and output value set U, such that the map

$$\begin{bmatrix} F+GDH & GC \\ B_2H + B_1DH & A + B_1C \end{bmatrix} : X \oplus X^{(r)} \to X \oplus X^{(r)} \qquad (3.14)$$

is stable. (Here $B := (B_1, B_2)$, and D acts only on the Y component.) The system Σ is *regulable* if it admits such a regulator.

Lemma 3.15: The system Σ is regulable if and only if it is asycontrollable and detectable.

Proof: For the sufficiency part; see Emre [10]. We prove here only the (easier) necessary part. In terms of the discrete-time interpretation, the map in (3.14) is just the closed-loop dynamics of the interconnection of Σ and $\Sigma^{(r)}$.

$$x(t+1) = Fx(t) + Gu(t), \qquad y(t) = Hx(t), \qquad (3.16a)$$

$$x^{(r)}(t+1) = Ax^{(r)}(t) + B_1u(t) + B_2y(t),$$
$$u(t) = Cx^{(r)}(t) + Dy(t). \qquad (3.16b)$$

Let x^* be in X, and pick any $x^{(r)*}$ in Z. Since (3.14) is stable, the sequence $(x(t), x^{(r)}(t))$ obtained solving (3.16) is stable. Thus the corresponding $u(\cdot)$ is a linear combination of stable sequences, and so is stable itself. It follows that x^* is a.c.; thus Σ is asycontrollable, and a dual argument establishes detectability. So Σ is regulable. #

IV. Stable Factorizations

We study here right factorizations $W = PQ^{-1}$ of a transfer matrix into stable matrices P, Q which satisfy the Bezout condition $AP + BQ = I$ with A, B stable. If q is a common stable denominator for the entries of P and Q, then also $W = YZ^{-1}$ and $CY + DZ = I$, where $Y = qP$, $Z = qB$, $C = q^{-1}A$, and $D = q^{-1}B$. Thus, when studying stable factorizations one may take P and Q to be polynomial without loss of generality.

Assume that $W = W(\Sigma)$ for a strictly causal detectable Σ. By the dual of Proposition 3.8, there exist stable linear maps M, N as in (3.12). If G would be the identity, then

$$W = H(zI - F)^{-1} \tag{4.1}$$

and (3.12) shows that (4.1) is a stable factorization of W. If G is present, but Σ is *reachable*, one may expect to be able to somehow eliminate G, since every state is reachable. This intuitive idea motivates the main result in this section, which generalizes the result in Hautus and Sontag [14, Theorem 6.11] as follows.

Theorem 4.2: A strictly proper transfer matrix W admits a stable right factorization if and only if it admits a reachable and detectable (strictly causal) realization.

Proof: Let $\Sigma = (X, F, G, H)$ be reachable and detectable, with $W = W(\Sigma)$. Since X is projective, there is for $n = n(X)$ a map

$$H_1 : X \to R^n \tag{4.3}$$

with

$$LH_1 = I = \text{identity in } X \tag{4.4}$$

for some $L: R^n \to X$. The reverse composition $H_1 L$ is just the projection on X, which will be thought of as a submodule of R^n. Consider the transfer matrix induced by $\Sigma_1 := (X, F, G, H_1)$:

$$W_1 := H_1(zI - F)^{-1}G. \tag{4.5}$$

Since Σ_1 is a split system, there exist by Theorem 2.11 polynomial matrices T, Q, C, and D, with Q admissible and

$$W_1 = TQ^{-1}, \tag{4.6}$$

$$CT + DQ = I. \tag{4.7}$$

Note that the image of W_1 (as a rational matrix) is included in $X((z^{-1}))$ since H_1 has image in X. Thus the image of T is also included in $X((z^{-1}))$, and so

$$T = H_1 LT. \tag{4.8}$$

Since $H_1(zI - F)^{-1}G = TQ^{-1}$, it follows by (4.4) that

$$(zI - F)^{-1}G = LTQ^{-1}, \text{ or } GQ = (zI - F)LT. \tag{4.9}$$

Let

$$P := HLT, \tag{4.10}$$

so that

$$W = HLW_1 = PQ^{-1}. \tag{4.11}$$

This will be the desired factorization (note that P and Q are, in fact, polynomial). We will now show the Bezout property. Since Σ is detectable, there are linear maps M, N as in (3.12). Composing with CH_1 on the left and with LT on the right, and applying (4.7), (4.8), and (4.9) gives

$$CH_1 M(zI - F)LT + CH_1 NHLT = CH_1 LT, \tag{4.12}$$

$$CH_1 MGQ + CH_1 NP = CH_1 LT, \tag{4.13}$$

$$CH_1 MGQ + CH_1 NP = CT = I - DQ. \tag{4.14}$$

So

$$AP + BQ = I \tag{4.15}$$

where we denote

$$A := CH_1 N \quad \text{and} \quad B := CH_1 MG + D. \tag{4.16}$$

Both A and B are stable because C, D are polynomial and M, N are stable.

Conversely, assume that $W = PQ^{-1}$ with P, Q polynomial, and that (4.15) holds for some stable A, B. Consider the "Q realization" corresponding to this factorization of W, as in Khargonekar [20, Theorem 4.4]. This is a reachable realization $\Sigma = (X, F, G, H)$, and all we need to establish is that Σ is detectable. We again introduce H_1, etc., as in (4.3)–(4.5), and define

$$R := W_1 Q = H_1(zI - F)^{-1}GQ. \tag{4.17}$$

Since the columns of Q are in the kernel (in fact, form a basis) of the input-to-state map associated with Σ_1, it follows that R is, in fact, a *polynomial* matrix. By reachability of Σ, there exist polynomial maps Y, Z with

$$(zI - F)Y + GZ = I = \text{identity in } X. \tag{4.18}$$

Thus

$$(zI - F)^{-1}GZ = (zI - F)^{-1} - Y. \tag{4.19}$$

Composing (4.15) with R to the left and with Q^{-1} to the right, and using (4.17) and (4.9) gives

$$RAPQ^{-1} + RB = RQ^{-1}, \tag{4.20}$$

$$RAW + RB = W_1, \tag{4.21}$$

$$RAH(zI - F)^{-1}GZ + RBZ = H_1(zI - F)^{-1}GZ, \tag{4.22}$$

$$RAH\left[(zI - F)^{-1} - Y\right] + RBZ = H_1\left[(zI - F)^{-1} - Y\right]. \tag{4.23}$$

Thus

$$M_1(zI - F) + N_1 H = H_1 \tag{4.24}$$

where

$$M_1 := RBZ + H_1 Y - RAHY, \qquad N_1 := RAH. \tag{4.25}$$

So $M(zI - F) + NH = I$ with $M := LM_1$ and $N := LN_1$. Note that M, N are stable because R, Y, Z are polynomial and A, B are stable. So Σ is detectable. #

By duality one also has the following.

Theorem 4.26: A strictly proper transfer matrix W has a stable left factorization if and only if it admits an asycontrollable and coreachable realization.

Neither Theorem 4.2 nor Theorem 4.26 is *a priori* self-dual: W may satisfy Theorem 4.2, but not Theorem 4.26. It is natural to ask about the existence of (only) *regulable* realizations, which is a self-dual condition. If either left or right factorizations exist for W, then Theorem 4.2 and Theorem 4.26 show that there are regulable realizations of W. The following result shows that (under weak extra assumptions) the existence of a regulable realization in fact *implies* that there are stable factorizations. The extra assumptions are on R and on the number of input or output channels. The (Krull) dimension of R (see, e.g., Gilmer [12]) is denoted by dim R; note that for polynomial rings or power series rings over a field, in r variables, one has $\dim R = r$. In Section VI it will be seen that for principal ideal domains R no constraints are necessary and, in fact, this holds for a slightly larger class of rings (see Remark 4.40). For more general R, we do not know if the assumptions on the number of input and output channels are really necessary. Our proof uses the fact that the module L in (4.30) is free, and this will, in general, be false.

Theorem 4.27: Assume that either $m=1$ or that $p=1$, or that R is Noetherian with $\dim R = r < \infty$ and $\max\{m, p\} > r+1$. Then the following properties are equivalent for a strictly proper transfer matrix W.

1) W has a stable left factorization;
2) W has a stable right factorization;
3) W has a regulable realization.

Proof: Note that Theorem 4.2 and Theorem 4.26 give that 1) and 2) imply 3). Since 3) is self-dual, it is then enough to prove that 3) implies 2). Let $\Sigma := (X, F, G, H)$ be an asycontrollable and detectable realization of W. We treat the case in which m satisfies the hypothesis; when p satisfies the hypothesis, the theorem can be proved using dual arguments. By asycontrollability, the map

$$[zI-F, -G]: S^{-1}(X \oplus U)[z] \to S^{-1}X[z] \tag{4.28}$$

is onto. Let L be its kernel, and denote $A := S^{-1}R[z]$. Note that $S^{-1}U[z] = A^m$. Thus

$$L \oplus S^{-1}X[z] \simeq A^m \oplus S^{-1}X[z]. \tag{4.29}$$

Using projectivity, we may assume that $S^{-1}X[z]$ is included in A^n, $n = n(X)$, and that there is some module Z with $S^{-1}X[z] \oplus Z = A^n$. So

$$L \oplus A^n \simeq A^{n+m}, \tag{4.30}$$

i.e., L has a free complement (is *stably free*) and has rank m. If $m=1$, L is free (see Lam [21, Theorem 4.11]). If R is Noetherian of finite dimension r, then $\dim R[z] = r+1$ (see, e.g., Gilmer [12, Theorem 30.5]) so $m > r+1 \geqslant \dim R[z] \geqslant \dim A$; thus one may apply Bass' cancellation theorem (see, e.g., Lam [21, Theorem 7.3]) to again conclude that L must be free. Let $\{v_1, \cdots, v_m\}$ be a basis of L. Without loss of generality, one may take the v_i polynomial over R. Consider the one-to-one map

$$A^m \to A^{n+m}: e_i \to v_i \tag{4.31}$$

whose image is included in the submodule $S^{-1}(X \oplus U)[z]$. Let

$$[v_1, \cdots, v_m] = \begin{bmatrix} T \\ Q \end{bmatrix} \tag{4.32a}$$

be its matrix. Note that

$$(zI-F)T - GQ = 0 \tag{4.32b}$$

since the v_i are in L. As the image of (4.31) is a direct summand of $S^{-1}(X \oplus U)[z]$, there exist stable maps Y, Z giving a left inverse, i.e., with $[Y, Z][v_1, \cdots, v_m] = I$, or

$$YT + ZQ = I. \tag{4.33}$$

We will now show that Q is admissible. Let $p = \det(zI-F)$; then $(zI-F)\operatorname{adj}(zI-F) = pI_n$, so the columns of

$$\begin{bmatrix} \operatorname{cof}(zI-F)G \\ pI \end{bmatrix} \tag{4.34}$$

are in L. It follows that the columns of (4.34) are in the span of the v_i, so

$$QR = pI \tag{4.35}$$

for some stable matrix R. Thus $(\det Q)\cdot(\det R) = p^n =$ monic, and Q is indeed admissible. Now let $P := HT$ (this is well-defined because the image of T is included in $S^{-1}X[z]$). Then

$$W = H(zI-F)^{-1}G = HTQ^{-1} = PQ^{-1}. \tag{4.36}$$

We claim that this is a stable factorization of W. Since the original system is detectable, there exist stable maps M, N such that

$$M(zI-F) + NH = I. \tag{4.37}$$

Multiplying on the right by $(zI-F)^{-1}GQ$ and on the left by Y,

$$YMGQ + YNP = YT = I - ZQ, \tag{4.38}$$

so

$$AP + BQ = I \tag{4.39}$$

for the stable matrices $A := YN$, $B := YMG + Z$. ∎

Remark 4.40: When $R[z]$ is a Noetherian integral domain of global dimension at most two, the ring $S^{-1}R[z]$ is projective-free, so that the module L in (4.30) is indeed free. Thus for principal-ideal domains (pid's) the above abstract argument is valid with no restrictions on m, p. The "concrete" constructions in Section VI also give the result, but they cannot be extended to non-pid's. We sketch here a proof, provided by W. Vasconcelos, of the projective-freeness claimed above. Take any projective $S^{-1}R[z]$-module P and consider it as a direct summand of a suitable free module $(S^{-1}R[z])^n$. Let M be the intersection in the latter of the sets $(R[z])^n$ and P, so that in particular $P = S^{-1}M$. Then M is a closed $R[z]$-submodule of $(R[z])^n$—because P is closed in the larger module—and is therefore reflexive. So M is a projective $R[z]$-module, by Bass' dimension-2 theorem. It follows that $T^{-1}M = M[(z)]$ is projective over

$R[(z)]$. But the latter is a ring of dimension one, so $T^{-1}M$ is, in fact, free over $R[(z)]$. It follows that M itself is free (Horrock's affine theorem), and so $P = S^{-1}M$ is also free, as required. #

V. Strictly Proper Factorizations

In this section we give a result on proper stable factorizations. Intuitively, the argument explained before Theorem 4.2 suggests that the existence of reachable and detectable realizations should imply that of proper stable factorizations, since M and N may be assumed both proper by (Proposition 3.8, statement 4). The problem is that one should not introduce extra delays when "eliminating" G as discussed in that argument. This difficulty can be overcome if one first finds a suitable stable bicausal isomorphism to premultiply the transfer matrix. The existence of such bicausal isomorphisms is known to be closely related to problems of "constant" state feedback, as discussed in the classical case by Hautus and Heymann [13]. Construction of constant stabilizing feedback laws for systems over rings is, in general, nontrivial (see, e.g., Kamen [17], Byrnes [3], and Bumby *et al.* [2]). For principal ideal domains (see Section VI) the result of Morse [23] ensures the existence of stabilizing K as needed in the result given here. This result connects our setup with the "proper stable" case in Desoer *et al.* [7]. We shall assume that R is an integral domain, and later note why the results also hold in general. A technical remark is needed first.

Lemma 5.1: Let X, Y be projective R-modules. Let the linear maps A: $Y \to Y$, B: $X \to Y$, C: $X \to Y$, and D: $X \to X$ satisfy $AB = CD$. Assume further that there exist linear maps M, N, P, and Q such that

1) $AM + CN = I =$ identity in Y, and
2) $PB + QD = I =$ identity in X.

Then $\det A = r \cdot \det D$, for some unit r.

Proof: Consider the linear maps

$$f := \begin{bmatrix} A & C \\ P & -Q \end{bmatrix} \text{ and } g := \begin{bmatrix} M & B \\ N & -D \end{bmatrix}: Y \oplus X \to Y \oplus X. \tag{5.2}$$

Then fg is triangular with identity diagonal, so f is invertible; say, $\det f = -r$. The conclusion then follows from the equality

$$f \cdot \begin{bmatrix} I & B \\ 0 & -D \end{bmatrix} = \begin{bmatrix} A & 0 \\ P & I \end{bmatrix}. \tag{5.3}$$

#

Theorem 5.4: Let $\Sigma = (X, F, G, H)$ be reachable and detectable. Assume that there is some K: $X \to U$ such that (the characteristic polynomial of) $F + GK$ is stable. Then $W(\Sigma)$ has a proper stable right factorization.

Proof: The first part of the proof repeats that of Theorem 4.2, up to (4.11). We use the same notations as are used there. Now let

$$Q^{\wedge} := Q - KLT. \tag{5.5}$$

Using (4.9),

$$(zI - F - GK)LT = (zI - F)LT - GKLT = GQ^{\wedge}. \tag{5.6}$$

Since $(X, F + GK, G, H)$ is again reachable, there is a right Bezout equation as in Lemma 5.1-i), with A, B, C, and D there equal to $zI - F - GK$, LT, G, and $Q^{\wedge}$, respectively. By (4.8) and (5.5) one also has an equation

$$(CH_1 + K)LT + DQ = I. \tag{5.7}$$

Thus Lemma 5.1 applies over $R[z]$ to give $\det Q^{\wedge} = r \cdot \det(zI - F - GK)$, r a unit in $R[z]$ (so in R). Thus $Q^{\wedge -1}$ is stable. Let

$$Q_1 := QQ^{\wedge -1}. \tag{5.8}$$

Since Q is a polynomial and $Q^{\wedge -1}$ is stable, it follows that Q_1 is stable. Also,

$$\begin{aligned} Q_1 &= ((Q - KLT)Q^{-1})^{-1} = (I - KLTQ^{-1})^{-1} \\ &= (I - K(zI - F)^{-1}G)^{-1}, \end{aligned} \tag{5.9}$$

so Q_1 is also proper. Let

$$P_1 := PQ^{\wedge -1}. \tag{5.10}$$

Then P_1 is stable and $W = P_1Q_1^{-1}$. Since $P_1 = WQ_1$, P_1 is also proper. It only remains to find a proper Bezout condition on P_1 and Q_1. By detectability, there are proper stable M, N as in (3.12). Composing on the right by $(zI - F)^{-1}GQ$ and using (4.8), we conclude that

$$MP + NGQ = LT. \tag{5.11}$$

Thus

$$\begin{aligned} Q^{\wedge} &= Q - KLT = Q - K(MP + NGQ) \\ &= (I - KNG)Q - KMP. \end{aligned} \tag{5.12}$$

Write $A := -KM$ and $B := I - KNG$; these are stable and proper. #

The same result is valid over nonintegral domains, but the above proof has to be modified slightly. This is because the determinants in Lemma 5.1 may not be *a priori* defined, for arbitrary projective X, Y over a ring with nonconnected spectrum. But the proof of (5.2) uses only the case $A =$ a characteristic polynomial and $D =$ a map between free modules, so that these are all well-defined up to associates.

The dual of Theorem 5.4 relates asycontrollable and coreachable realizations to the existence of proper stable *left* factorizations.

VI. The Case $R =$ Principal-Ideal Domain

The case in which R is a principal ideal domain is of special interest from a system-theoretic viewpoint, in particular, the ring $\mathbf{R}[\theta]$ of real coefficient polynomials is used in applications to delay differential systems, "2-D"—or "image processing" systems (note: transfer matrices are now rational in the two variables θ, z,) and single-parameter families of systems. Rings of rational functions with no real poles are used also for some of these applications, as well as rings of rational functions $\mathbf{R}[(\theta)]$. More generally, other rings of functions in one variable also appear sometimes, e.g., the ring of real-analytic functions in θ; such more general rings, although not necessarily pid's, still share many of the properties to be shown below. More precisely, for "elementary divisor rings" one can generalize the results in this section; see Bumby *et al.* [2] for the needed pole shifting results. Another pid of interest, the ring of integers, appears naturally in modeling fixed point digital implementations of systems.

In this section, R denotes a pid. Most statements can be considerably simplified in this case, and since the canonical realization $\Sigma(W)$ of a transfer function is necessarily projective (in fact, free), factorability properties can be checked directly on $\Sigma(W)$. We shall give first an elementary proof of Theorem 4.27 which does not involve any restrictions on m. The intuitive idea is straightforward: given a regulable system, restrict to its reachable subsystem. The nontrivial part is proving that this subsystem is indeed detectable. An easy lemma is needed.

Lemma 6.1: Let $\Sigma_1 = (F_1, G_1, H_1)$ be a factor system of $\Sigma = (F, G, H)$, i.e., $X = X_1 \oplus X_2$, and

$$F = \begin{bmatrix} F_1 & 0 \\ A & B \end{bmatrix}, \qquad G = \begin{bmatrix} G_1 \\ C \end{bmatrix}, \qquad H = (H_1, D). \tag{6.2}$$

If Σ is asycontrollable (respectively, detectable), then Σ_1 is also asycontrollable (respectively, detectable).

Proof: If Σ is asycontrollable, take stable rational M, N with

$$(zI - F)M + GN = I \ (= \text{identity in } X). \tag{6.3}$$

Let M_1 be the projection of M on X_1. Then

$$(zI - F_1)M_1 + GN = I \ (= \text{identity in } X_1) \tag{6.4}$$

and M_1 is again stable. This is similarly the case for detectability. #

Theorem 6.5: Let Σ be regulable. Then there is a system Σ_1, which is reachable and detectable, such that $W(\Sigma) = W(\Sigma_1)$.

Proof: We may assume that $\Sigma = (R^n, F, G, H)$ is weakly reachable, i.e., that the rank r of the span in (2.2) is equal to n. (This is because a weakly reachable subsystem is always a factor of the original system: take $X_1 :=$ kernel of any map T of rank $n - r$ which is zero on reachable states, and apply Lemma 6.1. Consider

$$\mathbf{g}(\Sigma) := [G, FG, \cdots, F^{n-1}G] \tag{6.6}$$

and let $D := \text{diag}(d_1, \cdots, d_n)$ be the Smith canonical form of $\mathbf{g}(\Sigma)$, where $d_i | d_{i+1}$ for $i = 1, \cdots, n-1$. Call $d_{ij} := d_i / d_j$ if $i > j$. By weak reachability, all d_i are nonzero. Using an appropriate T in $GL(R, n)$, we may assume that

$$\text{the columns of } D \text{ span the reachable states.} \tag{6.7}$$

Let $F = (f_{ij})$. We claim that

$$d_{ij} | f_{ij} \qquad \text{when } i > j. \tag{6.8}$$

To establish the claim, just apply F to $\text{diag}(0, \cdots, d_j, \cdots, 0)$. This must be again reachable, i.e., in the span of the columns of D. Comparing ith rows gives that $f_{ij} d_j =$ multiple of d_i, from which (6.8) follows.

Consider the system Σ_1 obtained by restriction to the reachable subset and a choice of basis for this subset. A concrete representation is $\Sigma_1 = (R^n, A, B, C)$, with $A = D^{-1}FD$, $B = D^{-1}G$, and $C = HD$. Let F_i (respectively, A_i) denote the submatrix obtained from the last i rows and columns of F (respectively, A). Since D is diagonal,

$$\det(zI - F_i) = \det(zI - A_i) \qquad \text{for all } i. \tag{6.9}$$

We claim that Σ_1 is detectable. Assume not. Then there is maximal ideal M of $S^{-1}R[z]$ such that

$$\text{rank}\begin{bmatrix} z^*I - A^* \\ C^* \end{bmatrix} < n \tag{6.10}$$

where the "*" is used to denote reduction modulo M. Thus there is a vector v over the residue field mod M such that $C^*v = (z^*I - A^*)v = 0$. Since $C^* = H^*D^*$, also $H^*(D^*v) = 0$. Since the original system is detectable, it follows that either

$$(z^*I - F^*)D^*v \neq 0 \tag{6.11}$$

or

$$D^*v = 0. \tag{6.12}$$

But (6.11) cannot hold because $D^*(z^*I - A^*)v = (z^*I - F^*)D^*v$. Thus (6.12) must hold. Let r be the smallest nonnegative integer such that $d_i^* = 0$ for all $i > r$. (Note that $d_i^* = 0$ implies that $d_j^* = 0$ for $j > i$.) Let s be the largest integer such that $v_i = 0$ for $i \leq s$. Since $D^*v = 0$, necessarily $r \leq s$. Since $d_j^* \neq 0$ for $j \leq r$, but $d_i^* = 0$ for $i > r$, it follows that $d_{ij}^* = 0$ for these (i, j). By (6.8), the corresponding f_{ij}^* also vanish. So the submatrix formed from the last $n - r$ rows and first r columns of F^* is zero. Similarly, the last $n - r$ rows of G^* are zero. By stabilizability of Σ,

$$\text{rank}[z^*I - F^*, G^*] = n, \tag{6.13}$$

so $\det(z^*I_{n-r} - F_{n-r}^*) \neq 0$. By (6.9) this holds also for A_{n-r}^*. But

$$(z^*I_{n-r} - A_{n-r}^*)w = 0 \tag{6.14}$$

where w is nonzero and consists of the last $n - r$ rows of v. This contradicts the nonzero determinant. Thus (6.12) cannot hold, and the theorem is established. #

The following summarizes all the results for the pid case.

Theorem 6.15: The following statements are equivalent for a strictly proper transfer matrix W:

1) W has a regulable realization;
2) W has a reachable and detectable realization;
3) the canonical realization $\Sigma(W)$ is detectable;
4) W has an asycontrollable and coreachable realization;
5) W has a stable right factorization;

6) W has a stable left factorization;
7) W has a stable proper right factorization;
8) W has a stable proper left factorization.

Proof: By Theorem 6.5 and its dual, 1) is equivalent to 2) and to 4). Since $\Sigma(W)$ is a factor of any reachable realization of W, Lemma 6.1 gives that 2) implies 3), which clearly implies 1). The equivalence of 2) and 5) [respectively, 4) and 6)] follows from Theorem 4.2 (respectively, Theorem 4.26). Finally, the equivalences of 7) and 8) with 5) and 6), respectively, follow from Theorem 5.4 and its dual. #

This theorem shows that factorability can be checked by first constructing a canonical realization and then checking detectability of this. For pid's there are various algorithms for canonical realization; see, for instance, Rouchaleau and Sontag [25] and Eising and Hautus [9]. Checking detectability is of varying difficulty, depending on the Hurwitz set considered. Two very simple examples are that of polynomial families of continuous-time systems, and that of delay-differential systems with arbitrary delay lengths. The latter studies existence of regulators for delay systems which are also described by delay systems and such that, for each value of the delay length, stabilization is achieved. We are grateful to E. W. Kamen for suggesting this example to us; as he conjectured, it is much easier to treat than the usual one (see a discussion in Hautus and Sontag [14] for detectability for a fixed delay length). In these examples, $R = \mathbf{R}[\theta]$.

In the delay case then, take the Hurwitz set

$$S := \{ p(\theta, z) \mid p \text{ monic in } z \text{ and } p \text{ has no zeroes in } L \} \tag{6.16}$$

where L is the set of complex pairs (s, v) with $\mathrm{Re}(s) \geq 0$ and $|v| \leq 1$ (see Kamen [18], [19]).

Lemma 6.17: For S as in (6.16), Σ is asycontrollable if and only if rank $[sI - F(v), G(v)] = n$ for all (s, v) in L.

Proof: In the terminology of Hautus and Sontag [14], we want to show that S is *perfect*, i.e., that the only maximal ideals of pr(R, S) are the obvious ones (evaluations at points of L). For this it is enough to prove that, given any pair (s^*, v^*) not in L, there is a p in S having it as a root.

If $\mathrm{Re}(s) < 0$, the polynomial $(z - s^*)(z - \bar{s}^*)$ (bar indicates conjugation) achieves the above purpose. If $|v^*| > 1$, then there is some integer n such that

$$k := (v^*)^{-n}(s^* + 1) \tag{6.18}$$

(and its conjugate) has magnitude less than 1. Let

$$p(z, \theta) := (z+1)^2 + |k|^2\theta^{2n} - 2\,\mathrm{Re}(k)(z+1)\theta^n \tag{6.19a}$$

$$= (z + 1 - k\theta^n)(z + 1 - \bar{k}\theta^n). \tag{6.19b}$$

By construction, $p(s^*, v^*) = 0$ (first factor in (6.19b) vanishes), while (6.19a) shows p has real coefficients. Assume p has a zero (s, v) in L. Then

$$|s + 1| = |k| \cdot |v^n|. \tag{6.20}$$

Since $|v| \leq 1$, $\mathrm{Re}(s) \leq 0$, a contradiction. #

For the family of systems example, we take S' as in (6.16), but we now use L' := set of pairs (s, v) with v real and with $\mathrm{Re}(s) \geq 0$.

Lemma 6.21: For S' as above, Σ is asycontrollable if and only if the condition in Lemma 6.17 holds over L'.

Proof: Again it suffices to find a suitable p passing through a given (s^*, v^*) not in L'. If $\mathrm{Re}(s) < 0$, we take the same polynomial as before. If v is not real, consider the real polynomial

$$b(x) := (x - v^*)(x - \bar{v}^*). \tag{6.22}$$

Note that there is a positive lower bound to the values of b on reals. Thus there is a real k with $kb(x) > \mathrm{Re}(s^*)$ for all real x. This implies that

$$\mathrm{Re}(s^* - kb(x)) < 0 \text{ and } \mathrm{Re}(\bar{s}^* - kb(x)) < 0 \tag{6.23}$$

for all real x. Let

$$p(z, \theta) := (z + kb(\theta) - s^*)(z + kb(\theta) - \bar{s}^*). \tag{6.24}$$

It is easy to see that this is a real polynomial having (s^*, v^*) as a root and [by (6.23)] without roots in L'. #

The above lemma says simply that asycontrollability (detectability) of the family $\Sigma(\theta)$ is equivalent to asycontrollability (detectability) for each individual system. The results on systems over rings then imply that constructions of compensators can be done polynomially in θ. This means that when the general structure of a system is known except for the precise value of the parameter, one may be able to design (off line) a compensator such that only tuning the parameter is needed when the original plant is completely identified.

We construct over this ring a very simple example in order to illustrate the various factorability conditions. Let a, b, c be three fixed real constants such that either $c \neq 0$ or $ab \neq 0$. Consider with $m = p = 2$ the transfer matrix W over $\mathbf{R}[\theta]$ with entries

$$w_{11} = (a + c\theta^2)/(s - 1) \tag{6.25a}$$

$$w_{22} = (s + 1 - \theta)/\left[s^2 - \theta s + (\theta - c\theta^2 - 1 - b)\right], \tag{6.25b}$$

$$w_{12} = w_{21} = 0. \tag{6.25c}$$

We use the Laplace variable "s" instead of "z" in order to emphasize here that we are viewing (6.25) as a family of continuous-time transfer matrices, parameterized by θ. The canonical realization has dimension 3 and is given up to isomorphism by the system

$$\dot{x}_1 = x_1 + u_1 \tag{6.26a}$$

$$\dot{x}_2 = x_2 + (b + c\theta^2)x_3 + u_2 \tag{6.26b}$$

$$\dot{x}_3 = x_2 + (\theta - 1)x_3 \tag{6.26c}$$

$$y_1 = (a + c\theta^2)x_1 \tag{6.26d}$$

$$y_2 = x_2. \tag{6.26e}$$

This is indeed canonical: the determinant of

$$[G, Fg_1] \tag{6.27}$$

is always $= 1$ (so the system is reachable), while the determinant of

$$[H', F'h_1'] \tag{6.28}$$

is $(a + c\theta^2)(b + c\theta^2)$, which is nonzero as a polynomial (giving observability).

To check the different factorability conditions for W, it is enough to study (6.26). Assume first that $ab \neq 0$, but $c = 0$. Then the determinant in (6.28) is a nonzero constant, so (6.26) is coreachable. Thus W splits for these values. Consider now the matrix

$$\begin{bmatrix} sI - F \\ H \end{bmatrix}. \tag{6.29}$$

Assume that $c = 1$, $b = 0$, and a is positive. The rows (4, 5, 2) of (6.29) give the minor $-\theta^2(a + \theta^2)$. The systems (6.26) are then all observable unless $\theta = 0$. But for $\theta = 0$ the system is nonetheless detectable, since rows (4, 5, 3) give a minor $a(s+1)$, which only has a root at $s = -1$. (Intuitively, from (6.26) it is clear that $\theta = 0$ destroys observability of x_1 from x_2, but in that case the x_1 coordinate asymptotically tracks x_2.) Thus the family is detectable as a system over the ring $\mathbf{R}[\theta]$ and the given Hurwitz set. One concludes that the original W has proper stable left and right factorizations. (Note that since we chose W symmetric, it is trivial in this case to obtain an asycontrollable and coreachable realization as claimed by the equivalence of 3) and 4) in Theorem 6.15: it is enough to take the transpose (dual) of the given realization.

Assume finally that the constants are $c = 1$, $b = 0$, and that a is nonpositive. Then the family is not detectable. Indeed, for some θ, the first column of (6.29) vanishes at $s = 1$ (i.e., x_1 is unstable and unobservable). So W has no possible Bezout factorizations of the types studied in this paper, and no realization of W can be made internally stable using a polynomially parameterized family of regulators.

REFERENCES

[1] N. Bourbaki, *Commutative Algebra*. New York: Addison-Wesley, 1972.

[2] R. Bumby, E. D. Sontag, H. J. Sussmann, and W. Vasconcelos, "Remarks on the pole shifting problem over rings," *J. Pure Appl. Algebra*, vol. 20, pp. 113–127, 1981.

[3] C. Byrnes, "Realization theory and quadratic optimal controllers for systems defined over Banach and Frechet algebras," in *Proc. IEEE Conf. Decision Contr.*, Albuquerque, NM, 1980, pp. 247–251.

[4] *Geometrical Methods for Linear System Theory*, C. Byrnes, and C. Martin, Eds. Dordrecht, The Netherlands: Reidel, 1980.

[5] *AMS-SIAM Symposia in Applied Math, Harvard 1979*, C. Byrnes and C. Martin, Eds. AMS-SIAM Publ., 1980.

[6] F. M. Callier and C. A. Desoer, "An algebra of transfer functions for distributed linear time invariant systems," *IEEE Trans. Circ. Syst.*, vol. CAS-25, pp. 651–662, 1978.

[7] C. A. Desoer, R. W. Liu, J. Murray, and R. Sacks, "Feedback system design: The fractional representation approach to analysis and synthesis," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 399–412, June 1980.

[8] S. Eilenberg, *Automata, Languages and Machines*, vol. A. New York: Academic, 1974.

[9] R. Eising and M. L. J. Hautus, "Realization algorithms for a system over a PID," *Math. Syst. Theory*, vol. 14, pp. 353–366, Oct. 1981.

[10] E. Emre, "On necessary and sufficient conditions for regulation of linear systems over rings," *SIAM J. Contr. Optimiz.*, vol. 20, Mar. 1982.

[11] E. Emre and P. Khargonekar, "Regulation of split linear systems over rings: Coefficient-assignment and observers," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 104–113, Feb. 1982.

[12] R. Gilmer, *Multiplicative Ideal Theory*. New York: Marcel Dekker, 1972.

[13] M. L. J. Hautus and M. Heymann, "Linear feedback: An algebraic approach," *SIAM J. Contr. Optimiz.*, vol. 16, pp. 83–105, 1978.

[14] M. L. J. Hautus and E. D. Sontag, "An approach to detectability and observers," in *AMS-SIAM Symposia in Applied Math, Harvard 1979*. AMS-SIAM Publ., 1980, pp. 99–136.

[15] E. W. Kamen, "Lectures on algebraic system theory: Linear systems over rings," NASA Contractor Rep. 3016, 1978.

[16] —, "An operator theory of linear functional differential equations," *J. Differential Equations*, vol. 27, pp. 247–297, 1978.

[17] —, "Linear discrete time systems defined over a commutative Banach algebra, with applications to two-dimensional systems," in *AMS-SIAM Symposium in Applied Math, Harvard, 1979*. AMS-SIAM Publ., 1980, pp. 225–238.

[18] —, "On the relationship between zero criteria for two variable polynomials and asymptotic stability of delay differential equations," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 90–91, 1980.

[19] —, "Linear systems with commensurate time delays: Stabilizability and stabilization independent of delay," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 367–375, Apr. 1982.

[20] P. Khargonekar, "On matrix fraction representations for linear systems over commutative rings," *SIAM J. Contr. Optimiz.*, vol. 20, Mar. 1982.

[21] T. Y. Lam, *Serre's Conjecture, Springer Lecture Notes in Math*, vol. 635. New York: Springer-Verlag, 1978.

[22] E. B. Lee and A. W. Olbrot, "On reachability over polynomial rings and a related genericity problem," in *Proc. Conf. Inform. Sci. Syst.*, Princeton, NJ, Mar. 1980, pp. 344–346.

[23] A. S. Morse, "Ring models for delay differential systems," *Automatica*, vol. 12, pp. 529–531.

[24] —, "System invariants under feedback and cascade control," in *LN Econ. Math. Systems 131*. New York: Springer-Verlag, 1976.

[25] Y. Rouchaleau and E. D. Sontag, "On the existence of minimal realizations of linear dynamical systems over Noetherian integral domains," *J. Comput. Syst. Sci.*, vol. 18, pp. 65–75, 1978.

[26] E. D. Sontag, "Linear systems over commutative rings: A survey," *Ricerche di Automatica*, vol. 7, pp. 1–34, 1976.

[27] —, "On split realizations of response maps over rings," *Inform. Contr.*, vol. 37, pp. 23–33, 1978.

[28] —, "Linear systems over commutative rings: A (partial) updated survey," in *Proc. IFAC '81*, Kyoto, Japan, 1981.

[29] M. Vidyasagar, H. Schneider, and B. A. Francis, "Algebraic and topological aspects of feedback stabilization," in *Proc. IEEE Conf. Decision Contr.*, Albuquerque, NM, 1980, pp. 891–895.

[30] B. Wyman, Stanford Univ., Stanford, CA, lecture notes, 1972.

[31] D. Youla, J. J. Bongiorno, and H. A. Jahr, "Modern Wiener–Hopf design of optimal controllers: Part I," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 3–15, 1976.

Pramod P. Khargonekar, for a photograph and biography, see p. 366 of the April 1982 issue of this TRANSACTIONS.

Eduardo D. Sontag was born in Buenos Aires, Argentina, on April 16, 1951. He received the Licenciado degree in mathematics from the University of Buenos Aires, and the Ph.D. degree in mathematics from the Center for Mathematical System Theory at the University of Florida, Gainesville.

From 1976 to 1977, he held a postdoctoral fellowship at the University of Florida, Gainesville. In 1977, he was appointed Assistant Professor of Mathematics at Rutgers, The State University, New Brunswick, NJ. He also worked in nonacademic computer-related positions. His major research interests lie in mathematical system theory, applied algebra, and theoretical computer science. He is the author of various papers in system theory and algebra, and has also written survey papers and a book on artificial intelligence. He has authored *Lecture Notes in Control and Information Sciences, 13: Polynomial Response Maps* (Berlin, Germany: Springer, 1979).

Dr. Sontag is a member of Phi Beta Kappa, of the American Mathematical Society, and of the Society for Industrial and Applied Mathematics.

ATE

LME